CHAPTER IV
A STUDY OF AIR POLLUTION-INDUCED CHRONIC ILLNESS

INTRODUCTION

At the time of the national awakening about environmental issues that occurred in the late 1960's, a great deal of public and scientific attention was focused on statistical relationships between air pollution and human health. While this research was undertaken with a large measure of academic curiosity, a major impetus was provided by Federal government agencies, such as the United States Environmental Protection Agency and its predecessors. The motivating factor for this agency encouragement was a laudable desire to establish scientific evidence for regulations designed to mitigate any detrimental health consequences of air pollution. For a time in the mid-1970's, the subject, though continuing to be discussed in scientific councils, did not capture much public attention, perhaps because of substantial reductions in the ambient concentrations of several common air pollutants. However, with the immediate threat that switching from oil and natural gas to coal fuels poses to the progress of a decade in controlling air pollution, the aforementioned statistical relationships are again a subject of public as well as scientific scrutiny.

In this paper, we assess the extent to which exiting epidemiological research can be interpreted as statistically demonstrating a relationship between air pollution and human health status. We also present some additional statistical research of our own. The next section is a critical review of the methodological underpinnings of existing research in air pollution epidemiology. So as not to exempt our previous work from this critical review, we devote a third section to self-appraisal. A fourth section presents some new empirical results meant to respond to several of the faults we confessed in the third section. The two concluding sections summarize what we think we have thus far learned and make some suggestions for future research.

A CRITICAL REVIEW OF OTHERS' WORK

Much of the recent work in air pollution epidemiology has focused upon

estimation of some version of the following expression:

$$H_i = a + bP_i + cX_i + u_i, \quad (1)$$

where H is a measure of morbidity or mortality, P is a measure of pollution, X is a set of other variables thought to influence health status, u is an error term that captures the effects of unmeasured influences upon health status, i indexes the individuals or groups of individuals in a sample, and a, b, and c are parameters to be estimated. Epidemiological work of this sort, a large part of which has been done be economists, presumes that there exists a distribution across individuals of tolerances to air pollutants and that there exist some individuals for whom any air pollution exposures whatsoever will trigger a decline in health status. This perspective may be contrasted with another, common to many epidemiological studies originating in the biomedical disciplines and sanctified in existing Federal clean air legislation, which posits a positive level of air pollution below which no individual will suffer a decline in health status[1].

Two recent empirical applications of the latter perspective are Morris, et al. (1976) and Bauhuys, et al. (1978). Inspired by the principles of experimental design, the researchers in each of these studies selected two communities similar in most respects other than air pollution. Using analysis of variance techniques, statistically significant differences in health status between the populations of the communities were then sought. Whether or not these differences were found, toxicological evidence from laboratory studies was then cited to provide a basis for rejecting or failing to reject air pollution as a cause of the difference. Many of the cited laboratory studies are, in principal, structured in the same fashion as the epidemiological studies; that is, the experimenter takes a treatment group and a control group of similar individual organisms and increases the pollution exposures of the treatment group until a decline in health status is observed. The pollution level at which this decline is first observed is then said to be the threshold at which pollution is universally unhealthy.[2]. Practitioners of this perspective generally agree that most substances commonly termed air pollutants can have deleterious human health effects. The controversies among them erupt over the threshold pollution levels at which these effects emerge and whether these threshold levels are found in everyday human environments. Because the methods provide no information on the magnitudes of any effects that do exist, the controversies are limited to questions on the statistical determination of the existence of an effect.

Unless all factors that contribute to differences in health status across individuals and locations can be controlled, the weaknesses inherent in empirical applications of the above perspective are apparent. In particular,

statistically significant differences between the health states of two groups of individuals may not be observable because the contributions of air pollution to the true differences are overwhelmed by uncontrolled factors. Any perceived threshold is then more a matter of experimental design rather than of effect: perception of 'where the threshold lies will differ with the extent to which the investigator is initially able to make his samples identical in all but their air pollution exposures. Moreover, even if the samples are identical, the outside observer gets the strong impression that there exists great confusion about the criteria for experimental design, the physiological and metabolic responses that constitute excess health impacts, the validity of extrapolating from animals to humans, and the processes that generate any defined health impact [3].

As is well known, the multivariate regression procedures usually used by economists investigating the health effects of air pollution allow explicit discrimination between the effects of air pollution, the effects of other observed control factors, and the effects of unobserved, presumably random factors. Although the estimated health effects of pollution will be biased if some of the assumed random factors vary systematically with pollution, the continuous covariation between health states and pollution that the procedures permit does not force one to adopt the ambiguous notion of a human health effects threshold before research is even initiated [4]. Neither is the investigator put in the uncomfortable position of having to assign the residual ("excess" deaths or illnesses) to something particular such as air pollution.

The first attempt to investigate the health effects of air pollution at a national level without the resumption of a threshold was the pathbreaking effort of Lave and Seskin (1970). Using 114 U.S. metropolitan areas as units of analysis, they employed single equation, ordinary-least-squares methods to regress 1960 mortality rates linearly upon ambient concentrations of sulfates and particulate, and other plausible influences upon mortality. They tentatively concluded that statistically significant health effects of air pollution existed. This original study has inspired a substantial number of similar subsequent studies, including the culminating effort of Lave and Seskin (1977) [5]. Without exception, all have discerned a close and substantial inverse association between mortality rates and one or more air pollutants. Recently however, two studies have become available that should give considerable pause to those wishing to accept the Lave-Seskin, et al. findings.

Smith (1977), using data for 50 U.S. metropolitan areas in 1968-1969, applied versions of the Ramsey (1969) tests for specification error in the general linear model to 36 different single equation specifications. These specifications were similar, and often identical, to those greeted with the

most approval by the authors of the Lave-Seskin, et al. literature. None of the specifications could pass all of the Ramsey (1969) tests at the 10 percent level, although four passed all tests except that for non-normal errors.

The Ramsey (1969) tests are meant to be used to assess conformity with the basic assumptions for error structure of the classical linear model. They give no hint about events when attempts are made to correct for one or more of the specifications errors. In a recent paper, Crocker-Schulze, et al. (1979, pp. 24-71) use 1970 mortality data from 60 cities while trying to correct for potential omitted independent variable and simultaneous equation problems. Upon adding measures of medical care, cigarette consumption, and diet to the single equation Lave-Seskin, et al. specifications, they found no statistically significant effect of nitrogen dioxide, total suspended particulate, and sulfur dioxide upon the rate of total mortality[6]. Retaining the former variables, and accounting for the plausible simultaneity between health status and medical care, did nothing to improve the statistical significance of the three air pollution variables. On the presumption that these findings were sufficient to demonstrate the weakness of the Lave-Seskin type results, the authors did not go on to account for the obvious simultaneity between median age (or percentage over 65 years) and mortality incidence, income and mortality incidence, and several other plausible sources of simultaneity.

The results obtained by Smith (1977) and Crocker-Schulze, et al. (1979) cast doubt upon the robustness of the Lave-Seskin, et al. estimates, in spite of the no-threshold perspective embodied in these estimates. Nevertheless, before dismissing the hypothesis of an inverse relation between everyday air pollution levels and health states, it must be recognized that Lave- Seskin, et al, may have been asking more of their data than it was capable of giving'. Less than one in every 100 people dies in the U.S. each year. No biomedical authority asserts that air pollution is the dominant cause of the deaths that do occur. Many take the view that it is the direct cause of no more than a small fraction of these deaths, although they would agree that it may be quite important in intensifying predispositions toward mortality. However, the general properties of the underlying processes that encourage this predisposition are ill-understood. Thus, even with quite large samples, available estimation techniques and a <u>priori</u> knowledge may be inadequate for distinguishing the mortality effects of air pollution in a human population sample from a host of similar and plausible minor contributing factors.

The possible inadequacy of many available techniques for estimating the existence and/or magnitude of air pollutant-induced mortality applies with special force, given the data Lave-Seskin and their successors had to employ. Their work can be interpreted as an attempt at establishing the probability of

a representative individual currently residing in a representative region dying in a given year from a geographically representative level of air pollution occurring in a representative year. Since they had no information about the distribution of influential health factors, including air pollution, across the urban areas constituting their units of analysis, the identifying variabilities of their samples were perhaps drastically reduced. [87] When this relatively low variability of the samples is coupled with what are probably substantial measurement errors in the air pollution variables, the baggage of additional explanatory variables and more sophisticated estimation techniques to correct for specification error that the data are able to carry must be rather light. The attempted corrections may serve only to misinform. Furthermore, that which is being corrected may be only an apparition since, as Crocker (1975, pp. 350-351) demonstrates, the measure of (the probability of) death, employing some group of individuals as the fundamental unit of observation, can differ from one group to another; there could be as many unique measures employed as there are groups.

The preceding remarks lead us to three conclusions. First, given the biomedical and economic subleties inherent in comprehending the etiologies of air pollution-induced mortality and morbidity, the estimates obtained from aggregated data used in the great bulk of extant studies are unlikely ever to be sufficiently compelling to establish a consensus. Only the use of actual individuals as fundamental units of observation is likely to provide enough strength in the data base to carry the requisite statistical burdens. Second, the statistical burdens that have to be carried might be considerably lightened if research concentrates on morbidity rather than mortality. The frequency, and most likely the identifying variability, of the former is greater by a factor of fifteen or twenty. Finally, because one's health status is influenced by the choices one makes about lifestyles, environmental and occupational exposures to possible toxics, and other health-influencing factors, economics can provide a priori hypotheses and an analytical framework to lend additional structure to epidemiological investigations. The relationships with which observed real world outcomes are consistent can, therefore, be further narrowed.

A CRITICAL REVIEW OF OUR WORK

Crocker-Schulze, et al., (1979) embodies both mortality and morbidity studies. The mortality study had the essentially negative purpose of empirically demonstrating that the estimates derived in Lave-Seskin type studies are not at all robust. The morbidity study had the more positive purpose of investigating air pollution and human health status with a data set better able to bear added statistical burdens and to accept hypothesis testing about the impact of man's free will upon health status. In this section, we briefly

discuss several entirely correct ways in which the morbidity study is susceptible to injury. Strangely, although the study has been carefully pursued by many interested parties, few have hit it where their thrusts could not even begin to be countered without additional work on our part. Here, we present some of those thrusts.

Depending almost entirely upon ordinary-least-squares (OLS), the morbidity study estimated the effect of air pollution upon self-reported health status measured as length of time chronically ill and annual frequency of acute illnesses. Expressions linear in the original variables were estimated for several 400 person samples independently drawn from all household heads in the Panel Survey of Income Dynamics (PSID) [Survey Research Center (1972)] who had always lived in one state. Although some attention was devoted to NO_2, air pollution was generally measured as the annual 24-hour geometric mean of SO_2 and/or TSP in the head's county of residence for the year (196775) from which the sample was drawn. In addition to air pollution, measures of the intensity of the head's illness, his biological and social endowments, life-style, and work, home, and outdoor environments were, when available, included as explanatory variables. Air pollution contributed positively and significantly to both chronic and acute illnesses in the majority of the unpartitioned samples. Upon combining these dose-response estimates with a simple recursive labor supply formulation, the economic impact of air pollution-induced chronic illness upon labor productivity was estimated to exceed that of air pollution-induced acute illness by nearly a factor of 20.

These results encouraged us to proceed further, particularly with respect to investigating air pollution-induced chronic illness. The obvious initial further step was to correct some of the outstanding technical problems in our treatment of the dose-response functions estimated from the PSID data. [97] These problems fall into three general categories: (1) the definition of self reported health status; (2) the factors used to explain self-reported health status; and (3) the algorithm used to estimate self-reported health status.

The PSID data on the chronic illness health status of household heads consists only of responses to four questions stated in the following order:

1. Do you have a physical or nervous condition that limits the type of work you can do or the amount of work that you can do?

2. How much does it limit your work?

3. How long have you been limited in this way by your health?

4. Is it getting better, worse, or staying about the same?

In the case of the first question, persons were asked for a yes or no answer, while for the remaining three questions the response called for was categorical. The response to question #3 was used as the dependent variable in our earlier analysis. However, the responses to this question were recorded categorically with the uppermost category being bounded only by age. Moreover, this response was conditional upon the response to question #1 and possibly question #2. For these reasons, interpretation of the earlier chronic illness dose-response estimates required a string of assumptions that may or may not have been important to stated results. In any case, in order to assess the validity of the earlier results, it is preferable to remove any clouding that the assumptions may have introduced. The response to question #1 is unambiguous.

Even though the response to question #1 is unambiguous in terms of self reported health status, it need not represent the respondent's clinical health status. More specifically, individuals may not be alike in the way they determine whether or not they are chronically ill. Economic factors including type of job, access to disability benefits, and other measures of the opportunity costs of not working may be important to this determination. For example, consider two persons who are alike in every respect other than their hourly wage. The person with the lower of the two wage rates will have a lower opportunity cost of not working. He may be perfectly healthy but desire to work fewer hours and use illness as an excuse, or he may actually be sick more often than his higher income counterpart because he does not find it economically advantageous to be as healthy.

The preceding suggests that our earlier estimated chronic illness dose-response expressions might be biased because economic determinants of self reported health status were omitted. In addition to these economic determinants, other, more traditional life-style, biological endowment, medical care, and environmental determinants were omitted or imperfectly measured. For example, the earlier estimates included no information on job accident rates, and used cigarette expenditures as an index of cigarette consumption. These variable exclusions and imperfectly measured explanatory variables can bias the estimated contribution of air pollution to self-reported health status.

Finally, given the chronic illness health status variable employed in our earlier work, the use of an OLS estimation procedure could have been inappropriate for two reasons. First, self-reported health status might have been determined jointly with some explanatory variables (e.g., leisure exercise, cigarette smoking, and medical care) that were also choice variables. OLS estimates of the chronic illness dose-response expression would

then be biased and inconsistent. Second, the health status variable was recorded in a categorical rather than in a continuous fashion. This means that hetero-skedasticity could be present in the OLS-estimated chronic illness dose-response expressions with a consequent introduction of biases in the standard errors of the air pollution coefficients. As McKelvey and Zavoina (1975) show, the use of OLS procedures with categorical dependent variables can cause the relative impacts of certain variables to be severly underestimated.

SOME NEW, BUT LIMITED RESULTS

In this section, we present some new results which, insofar as available data allow, correct partially or wholly for the technical problems raised in the previous section. The outstanding failing of these new results is that we do not construct an explicit analytical model to account for the economic determinants of self-reported health status. Instead, we do no more than introduce explanatory variables such as family assets and union membership that would plausibly have a role to play in expressions derived from any analytical model dealing with the effect of the opportunity costs of not working upon perceived own health status.

Table 1 lists the variables we employ. Alcohol expenditures, numbers of daily cigarettes smoked, free access to medical care, physician population, carcinogenic potential in the workplace, precipitation, workplace job accident rate, current transfer income, and union membership all represent variables that did not appear in our previous chronic illness dose-response expressions. Separate structural expressions are estimated for numbers of daily cigarettes smoked, whether or not the individual has medical insurance, and whether or not he participates in strenuous leisure exercise on the presumption that they are jointly determined with health status. To account for plausible nonlinearities with respect to the impact of age and food expenditures on health status, squared, as well as original, values are entered for these variables.

In view of the categorical nature and the simultaneity of the dependent variable, the estimation technique selected was the two-stage limited dependent variables (2SLDV) approach suggested by Nelson and Olson (1978). More specifically, the estimation procedure these authors propose is to:

(i) Estimate the reduced form of the structural systemby applying an appropriate maximum likelihood technique to each.

(ii) Form instruments from the “predicted” values of the

TABLE 4.1

COMPLETE VARIABLE DEFINITIONS

Self-Reported Health Status Variables

DSAB - Limitation on work = 1; otherwise = 0

LDSA - Disabled for ≤ 2 years = 1; 2-4 years = 2; 5-7 years = 3; ≥ 8 years = 4; otherwise = 0.

Biological and Social Endowment Variables

AGE - Age in years.

EDUC - Completed 6-8 grades = 2; 9-11 = 3; 12 grades = 4; 12grades plus non-acedemic training = 5; college, no degree=6; college degree = 7; advanced or professional degree = 8; otherwise = 1.

FMSZ - Family size in number of persons in housing unit.

POOR - Stated that parents were poor ".. ..when you were growingup..." =1; otherwise = 0.

SEX - Male = 1; Female = 0.

Lifestyle Variables

ALKY - Annual alcohol expenditures X 10^2 per adult family member.

CIGN - Number of daily cigarette packs smoked per adult family member. This variable was calculated by dividing the PSID data on 1970 cigarette expenditures by the 1970 retail price of a pack of cigarettes in the 1970 state of residence. Retail price data was taken from Tobacco Tax Council, Inc. (1978, pp. 67-69).

FOOD - Family food consumption relative to food needs standard in percent. Consumption refers to food expenditures in dollars and includes amounts spent in the home, school, work, and restaurants, as well as the amount saved in dollars by eating at work or school, raising, canning, or freezing food, using food stamps, and receiving free food. The food needs standard is in dollars and is based on USDA Low Cost Plan estimates of weekly food costs as published in the March 1967 issue of the Family Economics Review. The standard itself-is calculated by multiplying the aforementioned weekly food needs by 52 and making a series of adjustments according to family size.

LEXR - Indication that dominant leisure-time activities involves strenuous exercise = 1; otherwise = 0. Strenuous activities were said to include fishing, bowling, tennis, camping, travel , hunting, dancing, motorcycling, etc.

Health Care Variables

HVET - Free access to medical care as a veteran or through medicaid = 1; otherwise = 0.

INSR - Has hospital or medical insurance = 1; otherwise = 0.

PHYS - Physicians per 10,000 population in county of residence on July 1, 1975. This data was obtained from U.S. Bureau of the Census (1978, Table 2).

Environmental Variables

CANX - An index.of workplace "carcinogenic potential" by two-digit SIC code as presented in Hickey and Kearney (1977) and determined by dividing their Table 8 by their Table 7. We are aware that these authors insist that "... the magnitude of the derived carcinogenic potential is not suitable for any health hazard inference" (p. iii).

COLD - Mean annual January temperature in the 1970 county of residence in F" X 10. This data is from U.S. Bureau of the Census (.1978, Table 4).

PRCP - Mean annual precipitation in inches X 10^2 in the 1970 county of residence. This data is from U.S. Bureau of the Census (1978, Table 4).

JACCR - Number of disabling work injuries in 1970 by 2 and 3-digit SIC code for each million employee hours worked. The data is from Table 163 of Bureau of Labor Statistics (1972).

SULM - Annual 24-hour geometric mean sulfur dioxide micrograms per cubic meter as measured by the Gas Bubbler Pararosaniline-Sulfuric Acid Method. The data were obtained from the annual USEPA publication, Air Quality Data - Annual Statistics, and refer to a monitoring station in the 1970 county of residence.

TSPM - Annual 24-hour geometric mean total suspended particulate in micrograms per cubic meter as measured by the Hi-Vol Gravimetric Method. The data were obtained from the annual USEPA publication, Air Quality Data - Annual Statistics, and refer to a monitoring station in the 1970 county of residence.

Pecuniary Variables

ASSETS - Sum of 1970 income in dollars X 10^2 from social security, retirement pay, pensions, annuities, dividends, interest, and rent.

UNION - Member of a labor union = 1; otherwise = O.

dependent variables using the observations from the sample on the exogenous variables together with the estimated reduced from coefficients obtained in the first step.

(iii) Replace the jointly dependent variables on the righthand side of the equations in the structural system with their instruments constructed in the second step.

(iv) Estimate the resulting relations by an appropriate maximum likelihood method.

As can be easily seen, this estimation procedure applied to a system of simultaneous equations is just two-stage least squares in the case where all jointly dependent variables are continuous over the entire real line. However, the approach of Nelson and Olson (1978) takes account of the fact that some dependent variables, particularly the DSAB variable of interest here, do not exhibit this type of behavior. They therefore suggest that an appropriate limited dependent variable technique be used in the estimation of both the reduced form and the structural form of the model. In this case, since DSAB is defined to take on only the values of zero or one, the probit model would appear to be the most appropriate of the alternative limited dependent variable methods.

The procedures outlined above were applied to a sample of 309 individual household heads drawn from the 1970 calendar year of the PSID sample. All individuals had always resided in the 1970 state of residence. We are, thus, able to control partially for the air pollution exposure history of the individual, given that relative 1970 pollution concentrations across residential locations are similar to the history of relative concentrations, The year 1970 was selected for detailed empirical analysis because the chronic illness dose-response expressions estimated for this year in Crocker-Schulze, et al. (pp. 105-109) were considered to be the best representatives of all the expressions for assorted years estimated by ordinary-least-squares from the PSID data.

The 309 individuals of the sample represent all individuals in the 1970 PSID calendar year data for whom we were able to obtain observations on each explanatory variable, including total suspended particulate and sulfur dioxide. It should be noted that this sample is unlikely to correspond to a random sample of the U.S. population. If anything, as a glance at the arithmetic mean values of the explanatory variables presented in Table 2 shows, the sample appears to include a somewhat disproportionately high number of female household heads, “poor” childhood backgrounds, and relatively low pecuniary values of family assets. For our present purposes, of course, a

random sample is unnecessary, given that the sample was not selected on the basis of whether or not the individual reported he suffered from a chronic illness.

The results of estimating the augmented (relative to our previous work) chronic illness dose-response' expression by the multivariate Probit estimator are reported in the last two columns of Table 2. As Poirier and Melino (1978) have shown, the coefficients are proportional to the change in the probability that an individual will report being chronically ill for a one unit change in the explanatory variable. Thus, for example, a male, is nearly twice as likely to report being chronically ill as is a female. Our use of the Probit estimater presumes that each individual has a threshold level of the explanatory variable below which he will not view himself as being made chronically ill. However, the estimator also presumes that there exists a transformation causing these threshold values to be normally distributed over our sample and, therefore, that there exist some individuals for whom even . minor levels of air pollution will cause them to report being chronically ill. The constant term is simply a shifter.

With the exceptions of CIGN, LEXR, and POOR, the signs of all coefficients coincide with a priori expectations. The combinations of signs for the AGE variables and the FOOD variables are consistent with increased likelihoods of reporting chronic illness at the extremes of age and diet adequacy with a reduced likelihood in the middle ranges. Increases in alcohol consumption, exposures to carcinogenic substances, accident risks in the workplace, physicians to originate or confirm the individual's self-diagnosis, and air pollution in the form of sulfur dioxide all serve to increase the chances of self-reported chronic illness. The coefficients of CANX and JACCR are probably biased downward, since they refer only to the current workplace, rather than to the individual's workplace history. On the other hand, consistent with the work of Tromp (1962) and others, high precipitation and low midwinter temperatures are less likely to make the individual feel chronically ill. Those variables such as ASSETS and UNION, representing factors thought to reduce the opportunity costs of feeling chronically ill, all contribute positively to the probability of reporting chronic illness. Similarly, more education and larger family size, variables which capture factors tending to increase the opportunity costs of feeling chronically ill, each have negative signs attached. Since people who are veterans and have medical insurance face lower marginal prices for medical care, they can be expected to consume more medical care thereby reduce the frequency of their chronic illnesses. The negative signs attached to HVET and INSR are consistent with this interpretation. Note that the coefficient attached to the latter variable is estimated from a system that accounts for the simultaneity between the likelihood of possessing medical insurance and the

TABLE 4.2

MAXIMUM LIKELIHOOD ESTIMATES OF SELF-REPORTED CHRONIC ILLNESS (DSAB)

Variable		Mean	Coefficient	Standard Error
AGE		39.36	0.084	0.054
$(AGE)^* \times 10^{-1}$		177.00	-0.776	0.582
ALDY		1.11	0.169	0.100
ASSETS		2.68	0.001	0.001
CANX		18.77	0.006	0.021
$\widehat{CIGN}$	$\{CIGN =$	1.73		
	$\{\widehat{CIGN} =$	0.64	-0.527	0.190
COLD		37.86	-0.025	0.015
EDUC		3.76	-0.087	0.162
FMSZ		3.22	-0.005	0.056
FOOD		1.80	-0.499	0.470
$(FOOD)^2$		3.90	0.089	0.095
HVET		0.19	-0.472	0.400
$\widehat{INSR}$	$\{INSR =$	0.72		
	$\{\widehat{INSR} =$	0.80	-1.223	0.490
JACCR		33.17	0.003	0.005
$\widehat{LEXR}$	$\{LEXR =$	0.18		
	$\{\widehat{LEXR} =$	-1.13	0.115	0.454
PHYS		24.08	0.007	0.010
POOR		0.52	-0.503	0.290
PRCP		39.77	-0.043	0.017
SEX		0.57	0.927	0.556
SULM		18.37	0.011	0.010
UN 10N		0.19	0.422	0.398
Constant			1.090	1.807

(-2.0) times log of likelihood ratio 85.609; statistically significant at the one percent for the χ^2 distribution with 21 degrees of freedom.

Observations at Unity 77
Observations at Zero 232

NOTE: No levels of significance are indicated because the asymptotic properties of the standard errors for this sample are not known. A simulation experiment with the simultaneous probit estimator suggested to Nelson and Olson (1978, p. 702) that its standard errors could be biased upward by as much as a factor of 1.6.

presence of chronic illness. Note also, however, that the results for these variables explaining the "demand" for chronic illness have not been derived from an explicit analytical model. The above interpretation may therefore be unwarranted.

Interpretations for the signs of CIGN, LEXR, and PCOR are less readily provided. It is possible that no one of these variables is a reasonable measure of the effect we were trying to capture. For example, CIGN represents the estimated number of current cigarettes smoked per adult family member. There is no obvious connection between this measure and the smoking history of the individual whose health status is being inspected. It is, of course, possible that those who are already chronically ill increase their smoking because of the greater utility it might then afford. As for LEXR, it appears from its estimated mean value that the expression used to calculate it did not perform very well. In addition, the perception of what constitutes strenuous exercise can differ across individuals. Again, strenuous exercise might yield greater utility for those who are already chronically ill, so that they are more likely to participate in it than are healthy individuals. Similarly, the current perception of whether one's parents were poor may be more a measure of one's current real income status relative to the former status of one's parents rather than an absolute measure of the latter's former status. Thus, extending the Dusenberry (1949) hypothesis to an intergenerational context, it might be that greater relative current real income may engender a sense of security reducing the opportunity costs of being chronically ill. Alternatively, the explanation for the unexpected negative sign might simply be that a selection process operated in the past to eliminate those who were less well genetically endowed and who also had poor childhoods.

A rank-ordering of the explanatory variables from the most to the least statistically significant results in the following: CIGN, INSR, PROP, POOR, ALKY, SEX, COLD, AGE, $(AGE)^2$, HVET, FOOD, UNION, SUM, ASSETS, $(FOOD)^2$, PHYS, JACCR, EDUC, CANX, LEXR, AND FMSZ. Thus, at least for the sample represented in Table 2, air pollution, as measured by annual 24-hour geometric mean sulfur dioxide, is less robust statistically than the climate variables but more robust than the measures of occupational hazards. However, as indicated in the table, SULM would appear to be statistically insignificant at conventional levels. This general conclusion holds when another air pollution variable, annual 24-hour geometric mean suspended particulate, replaces the measure of sulfur dioxide used in Table 2. Upon doing this, a coefficient of 0.006 with a standard error of 0.007 is obtained. Given that the standard errors of the simultaneous probit estimator are thought to be biased upward (perhaps by as much as 1.6 according to Nelson and Olson (1978, p. 702), the actual effect of air pollution on self-reported health status may be more significant than our results indicate. Nevertheless, even if the standard error on the air

pollution coefficients are in fact biased upward by a factor of 1.6, the statistical significance of these coefficients remains questionable.

In order to provide another basis for comparison with Crocker-Schulze, et al. (1979), we substituted the measure used for the length of chronic illness (LDSA) in our earlier work for the dependent variable in Table 2. The system was estimated by the two-limit simultaneous probit technique employed in Nelson and Olson (1978). Again, the results obtained were not inconsistent with our previous OLS estimates. In fact, the magnitudes of the air pollution coefficients were almost twice those obtained in the OLS results. However, as Poirier and Melino (1978) demonstrate, the coefficients of an explanatory variable in a truncated regression procedure such as probit is proportional to, but not equal to, the partial derivative of the conditional mean of the dependent variable with respect to a one unit change in an explanatory variable. This factor of proportionality, which is identical for each coefficient in a regression, can be determined when the variance of the untruncated variable is known. For the PSID data set, this variance is unknown.

WHITHER FROM HERE

The motivation for this paper, as well as our previous work in the area, originated in our convictions that economic analysis and its empirical techniques could contribute to the resolution of certain recurring puzzles in studies of the incidence and severity of diseases in human populations, particularly the epidemiology of air pollution. We have viewed human health status as a decision variable and have therefore been able to employ economic theory as a means of providing more a priori structure for the analysis of epidemiological data. Considering only the empirical results reported in the previous section, it seems we have not yet provided enough information on structure for resolution. We have by no means, however, exploited all the conceivable economic-behavioral structural relations from which restrictions might be obtained.

One might introduce more statistical information by quasi-replication of the structures already estimated; that is, we could pull additional samples from the PSID data set and estimate for each of those samples the same two structures already discussed. This strategy has been used [Crocker-Schulze, et al. (1979)] in an earlier substantially less rigorous treatment of the same data.

Alternatively while retaining the structure that economic analysis and epidemiology provide, we can draw upon knowledge in biophysics, biochemistry, and bioenergetics to a much greater degree than previous studies in air

pollution epidemiology appear to have done. In a manner consistent with human capital theory, as some existing work has in fact already done [e.g., Cropper (1977) and Crocker-Schulze, et al. (1979)]. The individual might be construed as having an initial health endowment that, due to natural aging, depreciates exogenously over time. However, by his decisions about life-style and his occupational and environmental exposures, he can either slow or accelerate this natural depreciation. An integral part of these human capital treatments has been the representation of a production function in implicit form where some crude measure of health status is determined by rather arbitrary assortments of the aforementioned collection of life-style, occupational, and environmental variables. We suggest, at least insofar as empirical treatments are concerned, that one can specify this production function in much more detail while retaining the human capital framework for the individual's decision problem.

As an alternative to traditional toxicological research emphasis upon metabolities and metabolic pathways, the Second Task Force for Research Planning in Environmental Health Science (1977, Chapter 14) recommends that more effort be devoted to building upon existing knowledge of the structure and function of particular organ systems such as the respiratory and cardiovascular systems. Contrary to most of the arcane (to an economist) basic research on the fundamental chemical processes at work in various metabolic pathways, much of the work on the determinants of the individual's research of organ function appears to be readily translatable into mere displays of the fact that within limits the same quality of some simple measure of the health status of the organ system, such as the ventilation capacity of the lung, can be obtained from various combinations of inputs[10]. In many cases, the responses of the health indicator of the organ system to various stresses follow well- known physical laws having specific functional forms and even particular values attached to coefficients.[11].

When writing down the individual's decision problem with respect to health status, we may be able to structure the problem more tightly by building the aforementioned information on organ system responses directly into the constraint set. Rather than having an implicit production function in which the value of a "self-reported, highly aggregated measure of health status (e.g., whether or not the individual is chronically ill) is explained by a collection of intuitively reasonable variables, one can employ a description that precisely maps a limited and well-defined set of major influential factors into a continuous scaler measure of the health of an organ system.

SUMMARY AND CONCLUSIONS

The preceding pages are not without technical sin. In particular, with

out rigorously explaining from whence they cone, we have introduced variables that are supposed to represent the opportunity costs of reporting or failing to report ones self chronically ill. Otherwise, however, by employing a more robust estimation procedure, by redefining the chronic illness variable, and by introducing better measures of cigarette smoking, hazards and toxic exposures in the workplace, medical care, and climate, we have responded to several well-founded. criticisms of the morbidity results in Crocker-Schulze, et al. (1979). On the basis of those new tests, we see no reason to alter our previous interpretation of the effect of air pollution upon self-reported chronic illness.

REFERENCES

1 In accordance with the eloquent argument of Calabresi and Bobbit (1978), one might attribute the dominance of this perspective in public policy settings to the fictions erected by societies to segment markets that would otherwise require explicit judgments about the relative worths of individuals' lives. Calabresi and Bobbit (1978) argue that these fictions seine to soften intolerable societal stresses. The purpose they serve in a scientific setting is not obvious.

2 Alternatively, the laboratory studies try to specify the intervening processes causing an observed health effect.

3 Apart from these issues, the practice of applying laboratory results to everyday human environments is questionable. As Anderson and Crocker (1971, p. 146) note, so as to remove all sources of stress other than air pollution, all other factors influencing health in the laboratory tend to be set at biologically optimal levels. Given that these biologically optimal levels exceed those found in everyday environments, it follows from the law of variable proportions that air pollution-induced health effects in the laboratory will exceed those found in everyday environments.

4 It should be noted that many biomedical authorities strongly dispute the biological existence and the policy relevance of thresholds for most environmental contaminants. Authors such as Epstein (1974), Goldsmith and Friberg (1977) argue that any positive amount of pollution induces ill-health effects for some individuals and increases the probability of ill-health for everyone exposed.

5 Among the more notable examples are: McDonald and Schwing (1973); Liu and Yu (1976); Mendelssohn and Orcutt (1979); Gregor (1977) and Koshal and Koshal (1973).

6 However, particulate was statistically significant in an expression explaining pneumonia and influenza related deaths. Sulfur dioxide was

statistically significant in an expression for deaths attributed to early infant diseases. Nitrogen dioxide would have been statistically significant in heart disease if a slightly less severe level of acceptance had been adopted.

7 In order to get the data to "give" more, the authors of the Lave-Seskin type work have' usually tested with the same data set several different functional forms and combinations of explanatory variables. The objective frequently seems to have been the maximization of certain summary statistics (e.g., the coefficient of determination) having no basis in any *a priori* hypothesis. We are unaware that the pretest or selection procedures surveyed in Wallace (1977) and Judge, et al. (1980, Chap. II) have ever been employed during these manipulations. If these procedures are not employed, the properties of the classical least squares estimators these authors typically use can be substantially altered; that is, the customary interpretations cannot be attached to estimated coefficients and standard errors.

8 Ambient pollution concentrations for a single year at single (usually downtown) sites served as proxies for the lifetime exposure histories of entire regional populations. For a succinct treatment of the trade-off between corrections for specification error and identifying variability when measurement error is present in an independent variable of interest, see Griliches (1977, pp. 12-13). The addition of imperfectly measured explanatory variables to the expression being estimated will bias downward the coefficients of the air pollution variables.

9 For now, we much prefer to leave accounting issues about what the estimate mean in terms of national economic impacts to more adventuresome types.

10 See Kao (1972, Chap. 111 and IV) for readily understood treatments of the lung as a mechanical pump and as a gas exchanger.

11 Many of these responses have been established in animal rather than human studies. The validity of extrapolating results from the former to the latter is a major source of controversy in biomedical studies of pollution effects upon organ systems.

BIBLIOGRAPHY

Anderson, R.J., Jr., and T.D. Crocker, "The Economics of Air Pollution: A Literature Assessment," in P.B. Downing, cd., Air Pollution and the Social Sciences, New York: Praeger Publishers (1971), 133-166.

Bouhuys, A., G.J. Beck, and J.B. Schoenberg, "DO Present Levels of Air Pollution Outdoors Affect Respiratory Health?" Nature 276(Nov. 30, 1978), 466-471.

Bureau of Labor Statistics, Handbook of Labor Statistics, 1972, Bull. 1735, U.S. Department of Labor, Washington, D.C.: USGPO (1972).

Calabresi, G., and D. Bobbitt, Tragic Choices, New York: W.W. Norton (1978).

Crocker, T.D., "Cost Benefit Analysis of Cost-Benefit Analysis," in H.M. Pesken and E.P. Seskin, eds., Cost-Benefit Analysis and Water Pollution Policy, Washington, D.C.: The Urban Institute (1975), 341-360.

Crocker, T.D., W. Schulze, S. Ben-David, and A.V. Kneese, Experiments in Air Pollution Epidemiology, Washington, D.C.: USEPA Publication No. 60015-79-001a (1979).

Cropper, M.L., "Health, Investment in Health, and Occupational Choice," Journal of Political Economy 85(Dec. 1977), 1273-1294.

Duesenberry, J.S., Income, Saving, and the Theory of Consumer Behavior, Cambridge, Mass.: Harvard University Press (1949).

Engel, G.L., "The Need for a New Medical Model: A Challenge for Biomedicine," Science 195(Jan. 22, 1977), 129-136.

Epstein, S. "Environmental Determinants of Human Cancer," Cancer Research, 34(Oct. 1974), 2425-2435.

Goldsmith, J.R., and L.T. Fribert, "Effects of Air Pollution on Human Health,"

in A.C. Stem, cd., The Effects of Air Pollution, 3rd cd., New York: Academic Press (1977).

Gregor, J.J., Intra-Urban Mortality and Air Quality, Corvallis, Ore.: USEPA Pulbication No. 60015-77-009 (1977).

Griliches, A., "Estïmating the Returns to Schooling: Some Econometric Problems," Econometrics, 45(Jan. 1977), 1-21.

Hickey, J.L.S., and J.J. Kearney, Engineering Control Research and Development Plan for Carcinogenic Materials, Cincinnati, Ohio: U.S. Public Health Service under Contract No. 210-76-0147 (Sept. 1977).

Judge, G.G., W.E. Griffiths, R.C. Hill, and T. Lee, The Theory and Practice of Econometrics, New York: John Wiley and Sons (1980).

Kao, F.F., An Introduction to Respiratory Physiology, Amsterdam: Excerpts Medics (1972).

Koshal, R.K., and M. Koshal, "Environments and Urban Mortality: An Econometric Approach," Environmental Pollution 4(June 1973), 247-259.

Lave, L.B., and E.P. Seskin, "Air Pollution and Human Health," Science 169(August 21, 1970), 723-733.

Lave, L.B., and E.P. Seskin, Air Pollution and Human Health, Baltimore: Johns Hopkins University Press (1977).

Liu, B., and E. Yu, Physical and Economic Damage Functions for Air Pollutants by Receptor, Corvallis, Ore.: USEPA Publications No. 60015-76-011 (1976).

McDonald, G.C., and R.C. Schwing, "Instabilities of Regression Estimates Relating Air Pollution to Mortality," Technometrics 15(1973), 463-481.

McKelvey, R.D., and W. Zavoina, "A Statistical Model for the Analysis of Ordinal Level Dependent Variables," Journal of Mathematical Sociology 4(1975), 103-120.

Mendelssohn, R., and G. Orcutt, "An Empirical Analysis of Air Pollution Dose-Response Curves," Journal of Environmental Economics and Management 6(1979), 85-106.

Morris, S.C., M.A. Shapiro, and J.H. Wailer, "Adult Mortality in Two Communities with Widely Different Air Pollution Levels," Archives of

Environmental Health, 31(1976), 248-254.

Nelson, F., and L. Olson, "Specification and Estimation of a Simultaneous-Equation Model with Limited Dependent Variables," International Economic Review, 19(October 1978), 695-709.

Poirier, D.J. and A. Melino, "A Note on the Interpretation of Regression Coefficients within a Class of Truncated Distributions," Econometrics, 46(September 1978), 1207-1209.

Ramsey, J.B., "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," Journal of the Royal Statistical Society Series B, 31(1969), 350-371.

Second Task Force for Research Planning in Environmental Health Science, Human Health and the Environment: Some Research Needs, Washington, D.C.: USDHEW Publication No. NIH77-1277 (1977).

Smith. V.K., The Economic Consequences of Air Pollution, Cambridge, Mass.: Ballinger Publishing Co. (1977).

Survey Research Center, A Panel Study of Income Dynamics, Ann Arbor: Institute for Social Research, University of Michigan (1972).

Tobacco Tax Council, Inc., The Tax Burden on Tobacco, Washington, D.C.: Tobacco Tax Council, Inc. (1978).

Tromp, S.W., Medical Biometeorology, Amsterdam: Elsevier Publishing Co. (1962) .

U.S. Bureau of the Census, County and City Data Book, 1977, Washington, D.C.: U.S. Government Printing Office (1978).

Wallace, T.D., "Pretest Estimation in Regression: A Survey," American Journal of Agricultural Economics, 50(August 1977) 431-443.

Chapter V
MEASURING THE BENEFITS FROM REDUCED ACUTE MORBIDITY

INTRODUCTION

The predominant view in economics is that individuals are unaware of the health effects of air pollution and therefore do not take them into account in making decisions (Lave 1972). Given this view, the appropriate way to measure the morbidity benefits of a reduction in pollution is to estimate a damage function and then assign a dollar value to the predicted decrease in illness. This, together with any reduction in medical costs, is what an individual would pay for a decrease in pollution if he treated his health as exogenous.

Unfortunately, this approach is inconsistent with the view, widely held in health economics, that individuals can affect the time they spend ill by investing in preventive health care. Support for this view is provided by Michael Grossman (1972a, 1972b, and 1975) whose work indicates that individuals diet, exercise and purchase medical services to build up resistance to illness. These findings suggest that if persons in polluted areas perceive their resistance to illness decreasing they will try to compensate by exercising more, smoking less or getting more sleep. Conversely, an improvement in air quality should lead to a decrease in preventive health care, and the value of this must be added to the benefits of pollution control.

Human capital theory thus implies that the damage function approach, by ignoring the value of preventive health care, understates willingness to pay for a change in air quality. This conclusion, it should be emphasized, does not assume that individuals know precisely the medical effects of air pollution. All that is necessary for a person to try and compensate for the effects of pollution is that he feels worse when pollution increases.

This paper presents a simple model of preventive health care, similar to that of Grossman (1972a, 1972b), and uses the model to define what a person would pay for a change in air quality. The model assumes that one can build up resistance to acute illness by increasing his stock of health capital; however, health capital decays at a rate which depends on air pollution. For

acute illness, willingness to pay as derived from the model, is greater than the benefit estimate computed using the damage function approach. To illustrate the size of this discrepancy estimates of willingness to pay are computed using data from the Michigan Panel Study of Income Dynamics.

A MODEL OF INVESTMENT IN HEALTH

The essence of the human capital approach to health is that each individual is endowed with a stock of health capital, H, which measures his resistance to illness. This stock can be increased by combining time, TH_t, with purchased goods, M_t, to produce investment in health,

$$I_t = TH_t^{1-\zeta} M_t^{\zeta} E_{1t}^{\xi_1} \cdots E_{nt}^{\xi_n}. \qquad (1)$$

Outputs of equation (1) include exercise, rest and nourishment. These will be affected by factors such as the individual's knowledge of health, or the presence of a chronic disease ($E_{it}, \ldots, E_{nt}$ in equation (1)).

For simplicity suppose that investment in health exhibits constant returns to scale so that the marginal cost of investment is constant and independent of I_t. This is reflected in equation (2) which gives the marginal cost of investment, π_t, as a function of the price of purchased goods, PM_t, and wage, W_t,

$$\pi_t = W_t^{1-\zeta} PM_t^{\zeta} E_{1t}^{-\xi} \cdots E_{nt}^{-\xi_n}. \qquad (2)$$

Investment in health increases the individual's health stock, H_t, according to equation (3),

$$dH_t/dt = I_t - \delta_t H_t, \qquad (3)$$

Health capital also deteriorates at the proportional rate δ_t since resistance to illness would decline if no investments were made in health.

The main motive for investing in health is that health capital affects time spent ill, TL_t. For empirical work it is most appropriate to assume a threshold relationship between health capital and illness since a large number of persons (half of the Panel Study sample) report zero days of illness each year. A discontinuous relationship between H_t and TL_t, however, makes the solution to the individual's choice problem difficult. We therefore assume that the individual views the log of illness as a decreasing function of the log of health capital.

$$\ln TL_t = y - \alpha \ln H_t, \qquad \alpha > 0. \tag{4}$$

This implies that time spent ill can be made arbitrarily small, although not zero.

Equations (3) and (4) suggest that the model, while appropriate for acute illness, should not be applied to chronic illness. In (4) a reduction in the health stock increases time spent ill; however, being ill in one instant does not reduce the stock of health capital in the next. This is reasonable only if TL_t refers to acute illnesses such as colds and the flu.

To simplify the model and facilitate estimation of willingness to pay (4) is assumed to be the only motive for investing in health. This reduces health to a pure investment good and implies that the only effect of health on utility is through the budget constraint.

In this case the decision to invest in health can be separated from the decision to purchase other goods. First, a path of investment in health is chosen to maximize R, the present value of full income net of the cost of investment, then utility is maximized, given R. In the present model full income is the market value of the individual's healthy time. If Ω is the total time available at t then $h_t = \Omega - TL_t$ is the amount of healthy time available. The present value of full income net of the cost of investing in health may therefore be written

$$\int_0^T [(W_t h_t - \pi_t I_t) e^{-rt} dt, \tag{5}$$

where T is length of life. The individual's problem is to choose the path of investment which maximizes (5) subject to (3) and (4).

When the marginal cost of investment is constant the solution to this problem is simple: at each instant the individual chooses an optimal level of resistance, H_t^*, and then determines the amount to invest in health from (3).[17] The optimal health stock is determined by equating the value of the marginal product of health capital, $W_t \partial h_t / \partial H_t$, to its supply price,

$$W_t \frac{\partial h_t}{\partial H_t} = \pi_t \left(r + \delta - \frac{d\pi_t}{dt} \frac{1}{\pi_t}\right). \tag{6}$$

The latter consists of three parts: the interest foregone by investing π_t in

health rather than at the rate r, the depreciation cost, $\pi_t \delta_t$, since each unit of health immediately declines by an amount δ_t, and a capital gain which accrues if the cost of investment is changing! If π_t is rising at approximately the rate of interest then the right-hand-side of (6) reduces to $\pi_t \delta_t$.

Substituting from (4) the optimal health stock may be written

$$\ln H_t^* = \frac{1}{1+\alpha} (\beta + \ln W_t - \ln \pi_t - \ln \delta_t), \quad \beta = \gamma + \ln a, \tag{7}$$

while time spent ill is given by

$$\ln TL_t^* = y - \frac{\alpha}{1+\alpha} (\beta + \ln W_t - \ln \pi_t - \ln \delta_t). \tag{8}$$

There are several ways that pollution could enter this model. The observation that individuals are ill more often in polluted environments could mean that pollution enters the equation for time spent ill, (4), with a positive coefficient. This, however, implies that two individuals with the same health stock are not really equally healthy. Instead, it seems preferable to assume that pollution physically alters the state of a person's health. [2/] This can be accomplished by making the rate of decay of health capital a function of air pollution, P_t,

$$\delta_t = \delta \cdot e^{\delta t} P_t^{\psi} S_t^{\phi} . \tag{9}$$

Equation (9) also implies that the rate of decay of health varies with age and with other factors, St, such as stress or pollution on the job. [3/]

Adding equation (9) to the model means that it is more costly to build up resistance to illness in polluted environments, hence individuals in polluted areas will chose to maintain lower health stocks and will be ill more often than persons in cleaner areas. Proponents of the damage function approach might argue that this is unrealistic since individuals are unlikely to know the precise form of equation (9). All that is necessary, however, for an individual to choose a lower health stock is that he feels less healthy (perceives δ_t to be higher) when pollution increases. Knowing the precise relationshiptbetween δ_t and P_t is irrelevant in choosing H_t^*.

THE VALUE OF A CHANGE IN AIR POLLUTION

We now consider the value to an individual of a small reduction in pollution at time t. Since a change in P_t affects net income only at t the value of a small percentage change in P_t is defined as

$$-\frac{dR}{dP_t} P_t = \left(\frac{d\ lnTL_t}{d\ lnP_t} W_t TL_t + \frac{dI_t}{dP_t} \pi_t P_t\right) e^{-rt}. \qquad (10)$$

The first term on the right-hand-side of (10) is the value of the reduction in sick time caused by a reduction in pollution. This is unambiguously positive. The second term describes the change in investment costs caused by a change in pollution. Reducing pollution increases the optimal health stock which, from (3), increases I*. A reduction in P_t, however, also reduces δ_t which lowers the gross investment necessary to maintain a given health stock. For the functional forms above the net effect of these factors is positive, implying that a reduction in air pollution reduces resources devoted to preventive health care and thus increases willingness to pay,

$$-\frac{dR}{dP_t} P_t = \left(\frac{\alpha\psi}{1+\alpha} W_t TL_t + \frac{\alpha\psi}{1+\alpha} \pi_t \delta_t H^*\right) e^{-rt} = 2\frac{\alpha\psi}{1+\alpha} W_t TL_t e^{-rt}, \qquad (11)$$

If equation (10) is compared with the measure of benefits computed under the damage function approach it is clear that the latter understates willingness to pay. Following Lave and Seskin (1977) the damage function approach would measure the value of the reduction in sick time caused by a reduction in pollution, plus any change in medical costs. Since medical costs are negligible for acute illness, the damage function measure would equal the first term on the right-hand-side of, (10). The second term, which measures the decrease in resources devoted to preventive health care, would be ignored. To indicate the magnitude of this term and to give some idea of the morbidity costs of air pollution we present estimates of (10) based on data from the Michigan Panel Study of Income Dynamics.

ESTIMATION OF WILLINGNESS TO PAY

To compute willingness to pay requires an estimate of $\alpha\psi/(1+\alpha)$, the elasticity of sick time with respect to pollution. Equation (8) suggests that this can be obtained by regressing the log of sick time on the log of pollution and other variables which determine the optimal health stock. Since a large number of persons report zero days of illness each year the appropriate statistical formulation of the equation is a Tobit model,

$$\begin{aligned} lnTL_{it} &= \text{undefined} && \text{if } x'_{it}B + u_{it} \leq 0 \\ lnTL_{it} &= X'_{it}B + U_{it} && \text{if } X'_{it}B + u_{it} > 0 \end{aligned} \qquad (12)$$

where $x_t = (1\ lnPM_t\ lnE_{1t} \ldots lnE_{nt}\ lnP_t\ lnS_t\ lnW_t\ t)$

$$B' = \alpha(1+\alpha)^{-1}(\text{const.}\ 1-\zeta\ -\xi_1\ \ldots\ -\xi_n\ \psi\phi\ -(1-\zeta)\ \tilde{\delta}),$$

and $u_{it} \sim N(0,\sigma^2)$ for all t. Consistent estimates of (12) may be obtained by maximum likelihood.

Table 1 contains estimates of (12) for men between the ages of 18 and 45 from the Michigan Panel Study of Income Dynamics. The dependent variable is days lost from work due to illness, adjusted for differences in weeks worked. Independent variables, apart from the wage, either determine the rate of decay of health capital or affect the productivity of time invested in health.

Two features of the data should be noted. Since the dependent variable cannot be observed for persons too sick to work the estimates in Table 1 are subject to selection bias. This problem is not serious, however, since only 3% of the sample is unable to work for health reasons. Secondly, the data support a threshold model such as (12) since approximately half of the sample reports zero days of illness each year.

Before computing willingness to pay we comment briefly on the performance of the independent variables in Table 1. The first four variables measure factors which affect the rate of decay of health capital--air pollution, pollution at work, parents' income (which may affect δ_o) and race.[57] The first three of these consistently have the expected signs and are significant in six out of eight cases. Race, when significant, implies that being white increases the rate of decay of health capital. The second four variables affect the productivity of time spent investing in health. The presence of a chronic condition has a large negative impact on the productivity of time invested in health and is therefore positively related to sick time. Education, being married and being cautious should increase the prevention received for a given expenditure of resources and are in most cases negatively related to illness.

The chief anomaly in the health equations is the behavior of the wage. A high wage, by increasing the value of healthy time, should increase H^* and reduce TL_t. In Table 1 the wage is either insignificant of positively related to illness. This could be caused by two factors. In the Panel Study the wage is computed by dividing labor income by hours worked. This is not a good measure of the marginal wage unless an individual receives the same wage for each hour worked. Secondly, as Grossman (1972b) has argued, the wage may act as a proxy for deleterious consumption habits, e.g. , eating rich food, which increase the rate of decay of health capital.

We turn now to estimates of willingness to pay. In Table 1 pollution is measured' by the annual geometric mean of sulfur dioxide, which has been linked

TABLE 5.1

HEALTH EQUATIONS FOR MEN 18-45 YEARS OLD[a]

Independent Variable	Interview Year[b] 1970	1974	1976
Constant	"'3.5474	-1.2320	-0.5084
	(1.1253)	(0.9599)	(0.9014)
Ln(SO_2 Mean)	0.2879	0.3168	0.3189
	(0.2140)	(0.2076)	(0.1823)
Works in Manufacturing[c]		0.5001	0.4828
		(0.3659)	(0.3133)
Parents' Income	-0.1832	-0.1310	-0.0150
	(0.0936)	(0.1182)	(0.0953)
Race (1=White)	0.7318	0.3768	-0.2950
	(0.2697)	(0.4052)	(0.3084)
Has a Chronic Health Condition	1.1972	0.6515	0.9347
	(0.4582	(0.2862)	(0.2602'
Yrs. of Schooling	-0.1317	-0.1091	0.0496
	(0.0795	(0.1170)	(0.0508
Marital Status (1=Married)	-0.9678	0.9321	-0.6639
	(0.5098	(0.4550)	0.3828
Risk Aversion Index[d]	-0.3970		
	(0,0881		
Ln(Wage)	0,7492	-0.0899	0.1719
	(0.2873	(0.3553)	0.2813,
σ	2.1460	2.1586	2.1689
	(0.1824)	(0.2656)	(0.1931)
n			

[a]The dependent variable in each equation is the log of [work-loss days/(days worked + work-loss days)]x365. Standard errors appear beneath coefficients.

[b]Each interview year corresponds to the previous calendar year.

[c]Not available in 1970. [d]Not available in 1974, 1976.

Sources: All variables are from the Michigan Panel Study of Income Dynamics except SO_2 which is from the U.S. Environmental Protection Agency.

with acute illness in epidemiological studies. No other pollution variables are included since collinearity between pollutants leads to insignificant coefficients if several variables appear together. SO_2 should therefore be regarded as a pollution index and willingness to pay estimates viewed as indicators of the order of magnitude of willingness to pay. For the interview years 1970, 1974 and 1976 the mean of SO_2 is asymptotically significant at the .10 level or better' (one-tailed test); furthermore its coefficient is approximately 0.3 in each year, despite differences in the specification of the health equation.

Consider now the amount an' individual would pay for an x% reduction in pollution. According to (11) this amount is

$$2(x/loo)\ \frac{d\ln TL_t}{d\ln P_t}\ W_t\ TL_t\ . \qquad (13)$$

In equation (12) the elasticity of sick time with respect to pollution is equal to $\Phi(X'_{it} B/u)$, the probability of being ill, times the coefficient of the log of pollution. Since $\Phi(X'_{it} B/u)$ can be approximated by the fraction of the sample which is ill, $\Phi(X'_{it} B/\sigma) = 0.5$ in each year, implying that the elasticity of sick time with respect to pollution = 0.15. [5] The expected value of TL_t, calculated at the sample mean of X_i, is approximately 40 hours in each interview year. [6]

Equation (13) thus implies that the average person in the 1976 sample, who earned $6.00 per hour, would pay $7.20 annually for a 10% decrease in the mean of SO_2. The damage function approach, by contrast, would put the value of a 10% reduction in pollution at only $3.60. In a city with one million prime-aged men this would understate the value of a 10% reduction in air pollution by $3,600,000 annually. Ignoring adjustments to pollution, therefore, could sizeably understate the value of an improvement in air quality.

REFERENCES

1 For this solution to be valid the resulting value of I_t must lie between 0 and $\bar{I}$, the maximum I permitted at any t. (That $\bar{I}$ exists is guaranteed by the fact that Ω and non-labor income are finite.)

2 It is also true that air pollution affects productivity of time spent exercising; however, not all time invested in health is affected in this way. It therefore seems inappropriate to incorporate pollution in the production function for health.

3 In the paper δ_t is viewed as exogenous, hence the possibility of altering δ_t by moving or changing jobs is ignored.

4 Age, which should also affect the rate of decay of health, was dropped from the equation for lack of significance.

5 Evaluated at the sample mean of X_{it}, $\Phi(X'_{it}B/\sigma) = 0.57$ in 1970; 0.50 in 1974; and 0.53 in 1976.

6 $E(\ln TL_{it}) = X'_{it}B\Phi(X'_{it}B/a) + \sigma\phi(X'_{it}B/a)$. If this expression is evaluated at the sample mean of X_{it}, $E(TL_t)$ is, respectively, 46, 38, and 41 hours in 1970, 1974, and 1976.

BIBLIOGRAPHY

Grossman, M., "On the Concept of Health Capital and the Demand for Health," J. Polit. Econ., 80(1972a), 223-255.

Grossman, M., "The Demand for Health: A Theoretical and Empirical Investigation," National Bureau of Economic Research, New York, 1972b.

Grossman, M., "The Correlation Between Health and Schooling," In Household Production and Consumption, N.E. Terleckyl, cd., Columbia University Press: New York, 1975.

Lave, L.B. "Air Pollution Damage: Some Difficulties in Estimating the Value of Abatement," In Environmental Quality Analysis, A.V. Kneese and B.T. Bower, eds., Johns Hopkins University Press: Baltimore, 1972.

Lave, L.B. and E.P. Seskin, Air Pollution and Human Health, Johns Hopkins University Press: Balitmore, 1977.

U.S. Environmental Protection Agency, Air Quality Data Annual Statistics, Research Triangle Park, Selected Years.

University of Michigan Institute for Social Research, A Panel Study of Income Dynamics, Procedures and Tape Codes, Volumes 2, 4, 5, and 6, 1976.

CHAPTER VI

AIR POLLUTION AND DISEASE: AN EVALUATION OF THE NAS TWINS

INTRODUCTION

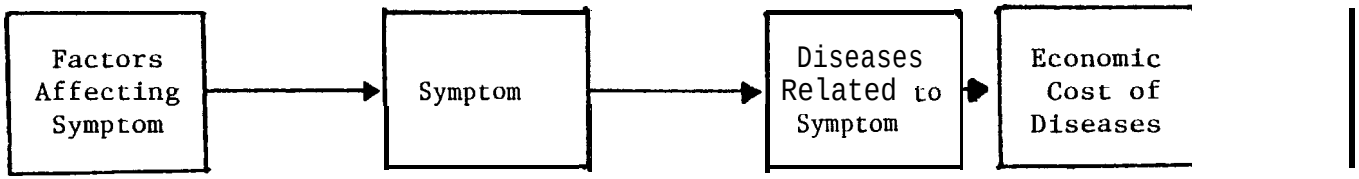
Human disease is caused by a mosaic of events, exposures, psychoses, genetic background, and the environment in which the individual resides. Air pollution is but one of the many factors potentially influencing morbidity and mortality rates of the population. The central question arises as to whether the net effect of air pollution can be assessed and measured such that a scientifically defensible estimate can be made of the change in health resulting from a change in ambient outdoor concentration of air pollutants. In recent years, a number of substantive studies have been undertaken to estimate this net effect. Lave and Seskin (1977) in their monumental work conclude that air pollution, when other factors are taken into account, contributes substantially to increased mortality across cities in the U.S. More recently, Graves and Krumm (1982) have demonstrated a connection (non-linear) between hospital admission rates and concentrations of carbon monoxide and sulfur oxides. Ostro has demonstrated a relationship between work loss days and particulate concentrations. Other studies have connected higher concentrations of air pollutants with indirect measures of lack of health [Gerking (1982).]

In this study we attempt to evaluate the impact of higher ambient concentrations of air pollutants on certain symptoms and reported diseases of a sample of approximately 14,000 twins who served in the Armed Forces during World War II. The simple idea underlying the study is that if there is a relationship between disease and air pollutant exposure, then exposure to higher concentrations of air pollutants, over time, should lead to a higher level of reported symptoms and incidence of certain diseases. Problems arise from many sources in this approach. For example, a symptom such as cough or shortness of breath can be related to the presence of many types of disease, or no disease at all. The presence of a cough, chest pain, and shortness of breath may be caused by asthma, emphysema, chronic bronchitis, or ischemic heart disease, among others. Secondly, the presence of a disease may not be detected because of a lack of one or more symptoms, or not seeking medical treatment. In addition, symptoms may be related to the presence of more than one type of disease. As one illustration, the individual may have both heart arrhythmia and emphysema, and yet exhibit shortness of breath as a single symptom. Finally, symptoms may not be accurately diagnosed and thereby reported on by the individual either because of a lack of basic medical understanding or other reasons. Also, there are substantial difficulties in relating symptoms to the prevalence of diseases, even though symptoms may emerge as a result of higher air pollutant exposures.

Factors other than the presence of air pollutants may have a significant effect on the occurrence of symptoms. Heavy smokers would tend to have a cough and perhaps shortness of breath regardless of air pollution concentrations. Air pollutants would then only exacerbate the presence of the symptom.

These and other qualifications must be kept in mind in evaluating the results reported later. A simple flow diagram (Figure 1) contains most of the hypotheses tested in this study. Examples of the factors proposed to influence the presence of symptoms are given in column 1. The list of symptoms recorded in the National Academy of Sciences twins data set are listed in column 2. A sample of the potential diseases that may be diagnosed from the symptoms are listed in column 3. Finally, in column 4 direct and indirect medical costs are given. In this study, primary efforts were made in relating factors affecting symptoms to symptoms and relating symptoms to the likelihood of a particular disease. As one example, increases in the level of total suspended particulate in the air may cause a greater number of individuals reporting severe chest pain (debilitating for more than one half hour) and shortness of breath when other factors such as cigarette consumption are taken into account. Severe chest pain over a period of time is one of the primary signals of the possibility of coronary heart attack or ischemic heart disease, although the signal may be for something else much less severe. Approximately 2 percent of individuals reporting severe chest pain have a coronary heart attack in the near future. Working through the chain of factors; symptoms, occurence of diseases, and economic cost of diseases, an estimate can be made of the impact of air pollutant exposure on economic costs. From some of the estimates reported later on, a 1 $\mu g/m^3$ increase in total suspended particulate concentration implies a $0.03 Per capita increase in economic costs associated with coronary heart attacks. However, these estimates should be viewed as purely experimental since many of the calculations and assumptions are new and have not been verified or replicated in independent analyses.

In the next section, a brief conceptual economic model is described where symptoms become a part of a household technology in solving medical problems. The following section contains a description of the data set. The next to last section contains the estimated regressions (one set) and final results on economic costs related to air pollutants.



Factors Affecting Symptom	Symptom	Diseases Related to Symptom	Economic Cost of Diseases
Diet	Cough	Chronic Bronchitis	Hospital Expenditures
Age	Shortness of Breath	Emphysema	Loss of Earnings
Asthma	Chest Pain	Bronchiectasis	Loss of Earnings Due to Death
Alcohol Consumption	Severe Chest Pain	Chronic Interstitial Pneumonia	Physicians Services
Income	Coronary Heart Attack	Ischemic Heart Disease	
Cigarette Consumption		Congenital Heart Disease	
Ambient Air Pollution		Cardiomyopathy	
		Cardiac Failure	

Figure 6.1 Major Relationships Examined and Statistically Estimated for the NAS Twins

MODEL DEVELOPMENT

A MODEL OF THE INDIVIDUAL'S HEALTH PROBLEM

It has been said by many people many times before that although they may not be rich, at least they have their health. This not only indicates the importance of one's health in the enjoyment of his life, but further suggests that an individual will normally have more than just a passive interest in the state or quality of his health. Stated in the terminology of the economist, one's health state is a valued good which yields utility to the individual.

There have been a reasonably large number of alternative economic models of health status proposed in the economic literature ranging from lifetime earnings concepts to labor market success. Most of these models concentrate on the effect of health status on the supply or productivity of labor (l). The general conclusion of these studies is that the occurence of diseases may reduce earnings by 20-30 percent through both amount of hours worked and the wage rate received. We have not discovered a study similar to this one which attempts to relate the incidence of disease, through symptoms, to specific causes, such as air pollution. Previous studies by the Wyoming group have focused on sorting out the demand and supply for medical services and how this is effected by air pollution (2). The issue of simultaneity in demand and supply is not addressed in this study.

It is safe to assume 'that an individual would like to have the best quality of health possible, but the procurement of such is not without costs. In particular, the individual may also gain utility from the consumption of goods which will adversely effect his health. For example, he may enjoy smoking cigarettes which has been linked to numerous lung ailments. Thus , the individual must balance his desire for smoking against his desire for good health. The acquisition of better health may also involve the necessary consumption of goods "which in and of themselves yield the individual dis-utility. For example, in order to increase the quality of his health state the individual may have to do some physical exercise when he prefers a more sedentary existence or he may have to eat types and quantities of food which are not to his liking (i.e., a salt-free diet or a simple weight-reducing diet) . Finally, the quest for good health may also involve more direct costs such as medical bills and possibly drugs such as aspirin, vitamins, insulin, or medicines to control blood pressure problems. Hence, one may envision the individual's problem with respect to his health as an economic one where choices must be made and tradeoffs considered between increased health quality and the costs of procuring it. In other words, within limits, an individual's health quality is a variable over which he possesses some control and which he will likely attempt to manage in some optimal fashion. It is the intent of this section to present a model of this problem and the relevant factors which are likely to influence the individual's choice. Particular emphasis will be placed on the role of air quality in this decisionmaking process.

The Utility Function

The utility function of an individual is a relationship between different quantities or bundles of goods and the satisfaction or happiness they provide to the individual in a specified time period. As noted above, the quality of one's health is likely to be a good which yields the individual utility. But numerous others could also be mentioned from French caviar to t-shirts. In this study, however, primary emphasis will be placed on those goods which are likely to either indirectly or directly effect the health of the individual. In particular, the individual's desires with respect to smoking, drinking of alcoholic beverages, nutrition, and the nature of his health state itself. Let the individual's utility function then be expressed as follows:

$$U_t = U_t(Q_t, C_t, B_t, E_t, H_t, X_t) \quad (1)$$

where:

Q_t refers to the air quality levels to which the individual is exposed at time t;
C_t is the quantity of cigarettes consumed at time t;
B_t is the quantity of alcoholic beverages consumed at time t;
E_t is the quantity of exercise (number of minutes) the individual engages in at time t;
H_t is the individual's perceived health status at time t;
X_t is the quantity of a composite good (i.e., all other goods) consumed at time t.

It appears reasonable to assume that the following relationships exist,

$$'_Q' \; U_H, U_X > 0; U_{QQ}, U_{HH}, U_{XX} < 0 . \quad (2)$$

With respect to the other variables, it is possible that either utility or disutility could be generated by the "goods" listed. If the goods are viewed as "goods" by the individual then the following relationships are likely to exist,

$$u_{C'} \; '_{B'} \; U_E > 0; U_{CC}, U_{BB}, U_{EE} < 0 . \quad (3)$$

If they are viewed as "bads" then,

$$'_C' \; '_B' \; U_E < 0; U_{CC}, U_{BB}, U_{EE} > 0 \quad (4)$$

of course, any combination of some of them as "goods" and some as "bads" would also be possible subject to the relationships relevant above.

Several points are relevant to this representation of the utility function. First, the state of one's health appears directly as a source of utility to the individual. It is likely that the health state actually is a joint "input" with the other goods in the "production" of utility but its importance in the utility function should nonetheless be downplayed any more than the role of energy inputs as joint inputs with agent inputs should in

the production of some output. Secondly, although the level of air quality may be viewed as a choice variable of the individual (he can effect it by living in different areas, for example), for the purposes of this investigation it will be taken as given and beyond the control of the individual in order to keep the number of adjustments the individual can make in response to it at a workable level. The inclusion of air quality in the utility function is a proxy for the aesthetic benefits the individual receives from the environment. As air quality deteoriates (i.e., visibility is reduced or the air begins to smell), it is likely that the individual will experience a loss of aesthetic benefits and so, a resulting loss of utility.

Finally, note that the individual may get utility from cigarette consumption which may adversely effect the utility he receives from the quality of his health. Thus, the tradeoff mentioned earlier and the need to more closely specify the nature of the effect on health.

The Respiration Process

In order to understand how various factors influence one's health state it is necessary to gain a rudimentary idea on how the human body works. The normal sequence of chemical changes in human calls depends on oxygen and hence, there exists the need for continuous supply. One of the chief end products of these chemical changes is carbon dioxide and hence, the need for continuous elimination of this waste. In simple single cell animals the intake of oxygen and the release of carbon dioxide occurs at the surface by diffusion. However, as organisms increase in size and complexity, a specialized structure is developed which functions to serve the needs of the various cells. In man this function, known as respiration, is performed by the respiratory system aided by the cardiovascular system.

Oxygen reaches the various cells in the body through three steps: (1) from the environment to the lungs, (2) the lungs to the blood stream, and (3) the blood stream to the cells. The movement of carbon dioxide out of the body is just in the opposite direction. Each of these steps may be discussed separately. The first step, referred to as ventilation, involves inspiration, or the breathing in of outside air and expiration, the breathing out of carbon dioxide. The driving physical force behind this process is Boyle's Gas Law which states that "volume varies inversely with pressure at a constant temperature."

On inspiration the primary muscle of the respiratory system, the diaphragm, pulls downward thus enlarging the cavity containing the lungs. This increase in volume, a la Boyle, causes a reduction in the pressure within this cavity with relative to normal "outside" pressures and so, causes air to rush in and expand the lungs as pressures are equalized. On expiration the diaphragm relaxes and just the opposite occurs forcing air out of the lungs. The substance of the lungs themselves is porous and spongy. Bronchial tubes (hollow air passageways) connect the lungs to the outside environment. Each lung is composed of a large number (billions) of air sacs called alveoli each covered by numerous capillaries. Thus, the ventilation process brings air into these alveoli on inspiration and removes air from them during expira-

tion. The makeup of the air inspired and that expired of course is not the same as that expired in percentage terms as it contains less oxygen (16 percent versus 21 percent) and more carbon dioxide than that inspired.

The second step in the respiration process is called external respiration and involves the passage of oxygen from the alveoli of the lungs to the blood stream (and vice versa, the passage of carbon dioxide from the blood stream into the alveoli). What occurs is the passage of oxygen through the alveoli membrane into the capillaries surrounding it and the opposite passage of carbon dioxide into the alveoli. This transfer occurs due to variances in partial pressures. As noted above, inspired air oxygen makes up a larger percentage of the total volume of air then it, does in the returning blood from the cells and so, has a higher partial pressure. Thus, as blood flows through the capillaries surrounding the alveoli, due to the pressure differentials, oxygen flows from the alveoli into the blood stream. Since the returning blood contains carbon dioxide released from the cells, the partial pressure differential is just opposite and so, carbon dioxide passes from the capillaries into the alveoli where the partial pressure of carbon dioxide is lower. This exchange is influenced by several factors: (1) the area of contact for the exchange, (2) the length of time blood and air are in contact (only about a second or two at any one time--at least once or twice a minute all the blood in the body passes through the capillaries of the lungs), (3) permeability of cells forming the capillary and alveolar membranes, (4) differences in concentrations of gases in alveolar air and the blood, and (5) rate at which chemical reaction takes place between the gases and the blood. Respiratory efficiency is also related to the number of red cells, hemoglobin content of these cells, and the area of the red cell (3).

The final step is internal respiration which involves the passage of oxygen from the blood into the tissue fluid and on into the cells and the reverse passage of carbon dioxide. After the exchange of oxygen and carbon dioxide in the lungs, the newly aerated blood (oxygen-carrying blood) is returned to the heart and then distributed to all parts of the body. As blood moves into the various capillaries, the partial pressure of the oxygen in it is high while that for carbon dioxide is low. Meanwhile, the reverse is true in the tissue fluid and cells since they have "used" previous supplies of oxygen and have created "waste" carbon dioxide. These pressure gradients once again result in the transfer of gases between the blood stream and the cells and thus, complete the respiration process.

The Oxygen Production Function

Given this somewhat brief description of what in reality is a most complex and not fully understood process, the human body, especially the respiratory and cardiovascular systems, may be viewed as a factory which processes an input (air in the environment) into a useful product for the cells of the body (oxygen). There is also the elimination of carbon dioxide, but this may be seen as just another side of the same coin. Considering useable and delivered oxygen to the cells as the output, an economic production function may be envisioned as follows,

$$02 = f(K, A) \quad (5)$$

where:

02 is the amount of oxygen delivered to various cells of the body during a specified time period

A is the total volume of environmental air of fixed quality, Q_t, which is inspired during the specified time period

K is the quality of the individual's "body capital" during the specified time period

In general, it is to be expected that

$$f_A, f_K > 0 \text{ while } f_{AA}, f_{KK} < 0 \quad (6)$$

but a closer examination yields even more information.

It should be clear that the two "inputs" in this production relationship serve different roles. The inspired air is material to be processed by the "body capital" (i.e., the various components of the human body--more on this below) into useable oxygen. Substitution across these two types of inputs may thus only be done up to a certain limit.* For example, if in a sedentary position an individual requires 20 liters of oxygen per hour then clearly at the very least the air inspired during an hour must contain 20 liters of oxygen (actually much more would normally be required since a relatively small percentage of the oxygen inspired is ever taken into the bloodstream). Thus, regardless of the state of the individual's body capital, a minimum of inspired air is required and cannot be substituted for. On the other hand, the body capital must be at some minimum level of efficiency in order to insure the 20 liters of oxygen eventually reaches the cells. So, for any given oxygen requirement during some period there are likely to exist minimum requirements of <u>both</u> inspired air and body capital quality and these requirements will increase with increased oxygen requirements. However, to the extent these minimums are attained some substitution between these inputs are possible. For example, one could achieve a given level of oxygen production in several manners. If the body capital is in a very poor state (but at least the minimum required) this may be offset by a higher flow of inspired air (increasing the rate of respiration). If the body capital is in fairly good shape, clearly less inspired air would be required. These relationships may be represented by the isoquant mapping of this production function shown in Figure 2.

Measured along the vertical axis is increasing body captial quality (measured in terms of some efficiency parameter), while increased quantities of inspired air <u>of a given</u> quality is measured along the horizontal axis. Each isoquant then represents those combinations of body capital quality and volumes of inspired air (again, of a given quality) which would yield a given amount of delivered oxygen to the cells, which as shown, is dependent on the activity level of the individual. Diminishing marginal rates of substitution are assumed. Note that each isoquant approaches both a vertical and horizontal asymptate to reflect the fact that for any level of oxygen produced there exist minimum requirements of both body capital and volumes of inspired air.

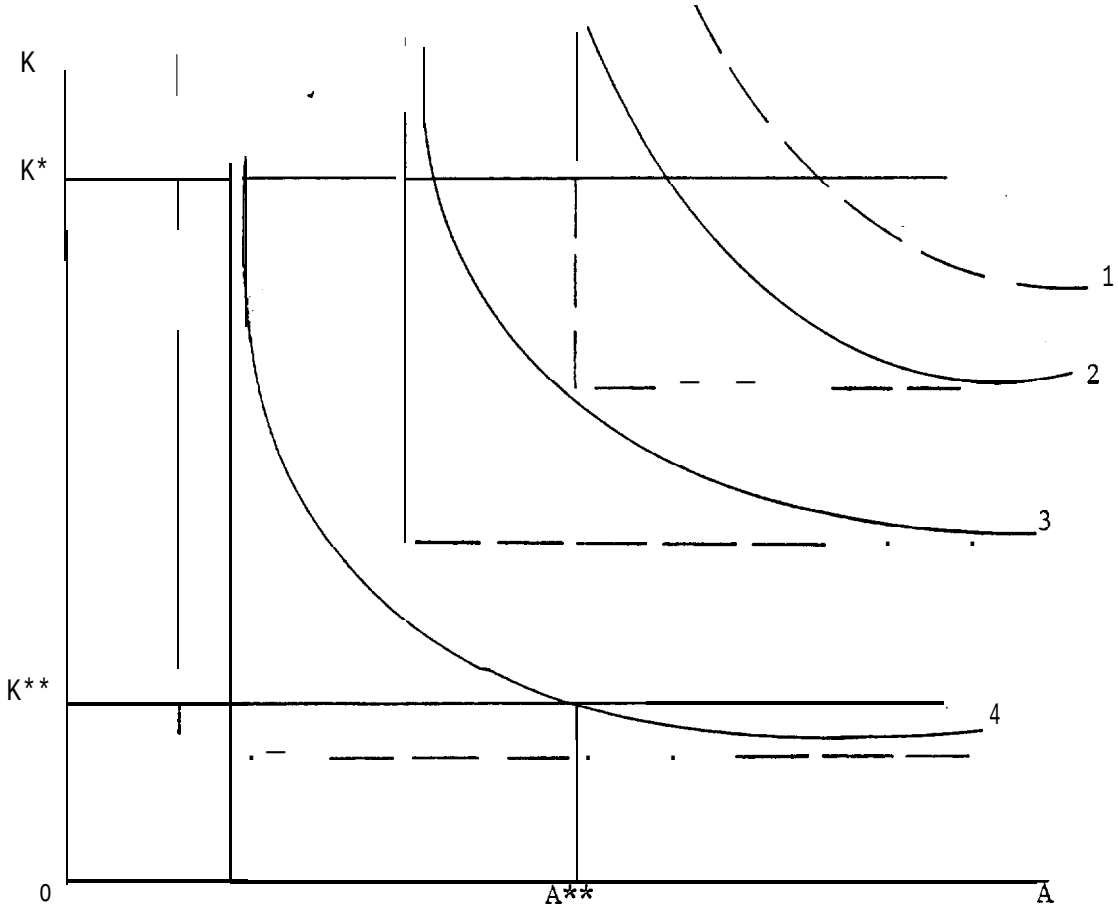


Figure 6.2 Conceptual Tradeoff Between Body Capital and Respiration

[1] O_2 required for heavy physical activity with an inferior air quality

[2] O_2 required for heavy physical activity

[3] O_2 required for light physical activity

[4] O_2 required for sedentary existence

This illustration of the "oxygen production function" of the human body will aid greatly in developing how an individual perceives the state of his health, however, let us digress at this point for a more indepth look at this variable called "body capital".

K or K_t represents the true health status of the individual as given by the quality of his "body capital," that is, the actual physical condition of his heart, lungs, and other components of his respiratory or cardio-vascular systems and the proficiency in which they perform their functions. Though not directly observable by the individual, in general one would expect that

$$K_t = K_t(K_o, \underset{\rightarrow}{Q_t}, \underset{\rightarrow}{C_t}, \underset{\rightarrow}{B_t}, \underset{\rightarrow}{E_t}, \underset{\rightarrow}{X_t}, \underset{\rightarrow}{M^K_t}) \tag{7}$$

where K. represents the individual's initial body capital quality endowment which would be based largely on inherited genes and the subscript $\rightarrow t$ refers to the full "time Profile" of consumption of the respective variable up to time t. This says that not only is the total consumption of some good, say cigarettes, C, important, but also the timing of this consumption. For example, given that an individual's body capital has some natural regenerative capabilities as many feel it does, than one would expect that someone who smoked one pack a day for a year 5 years ago might have a better state of body captial today than someone who smoked a pack a day for the last year. Thus, the quality of one's true health status is probably dependent on cumulative doses, as well as, the timing of those doses. This type of dependence is difficult to model, however, most relevant information may be captured by the following:

$$K_{t+1} - K_t = AK_t = g(K_t, Q_t, C_t, B_t, E_t, X_t, M^K_t) - \delta K_t \tag{8}$$

where K_t would include much of the information concerning past loadings of Q, C, etc. and δ represents a natural decaying factor of the quality of one's body capital with age. Generally it seems reasonable to assume the following,

$$g_Q, g_C, g_B < 0 \text{ and } g_E, g_{M^K} > 0 \tag{9}$$

given the latest medical evidence available (remember, the function g attempts to describe the actual change in one's true health status given a certain level of outside influences and that these true relationships are still not wholly determined by the medical profession). M^K denotes the amount of medical services and/or medicines purchased by the individual to improve the state of his health, i.e., vitamins, medicine to control blood pressure, or simply advice from a doctor. Since X is a "catch-all" including all other goods, it is uncertain how it will over time effect the level of K_t. Finally, included in the behavior of g would be some account for the natural regenerative capability of the body capital. In other words, for levels of Q, C, and B below some threshold level for each, one would expect g to be positive to reflect an improvement in body capital.

The Individual's Perceived Health Status

Given a level of K determined as in (7), let us return to Figure 2. Clearly, if K is at some level such as K* the individual should observe little problem with lack of oxygen. However, if his level of K were more like that of K** then note that light physical activity becomes impossible for him and even a sedentary existence requires more inspired air, A**, then the individual with K* quality (A*). This second individual will thus be getting a symptom (i.e., shortness of breath or chest pain if his heart must do extra duty to process more air) that something is wrong.

Another manner in which a symptom, a physical response of the body, might occur involves the level of air quality. However, suppose the air quality was worse. For a lower level of air quality it is likely that the isoquants of Figure 2 would shift in a northeasterly direction. That is, to produce a given amount of delivered oxygen would require both more inspired air (since the useable portion of this air would be less) and a higher quality of body capital since more of the material input would have to be processed. This suggests that an individual with a given level of K may experience no symptoms in a "good" air quality situation, but as air quality deteoriated symptoms would arise as the minimum requirements of inspired air rose.

Given the above, a symptom, an observable phenomenon to the individual, has basically two sources--a deteoriation of body capital or a deteoriation of air quality. With respect to air quality then it is possible to distinguish between its chronic effects (its effects on the quality of body capital) and its acute effects (its effects on changes in the useable nature of the material input--inspired air). So, the advent of a symptom may be the result of a true deteoriation of health status or simply the result of deteoriating environmental quality (wherein health status is actually not in jeopardy). Take coughing for example. This symptom could occur because the quality of body capital has been reduced to low levels and so even with good quality air the individual coughs (for example, the individual could be a long-time smoker and this has led to emphysema wherein many of the alveoli of the lungs have been rendered all but unusable). On the other hand, coughing could occur because of a high concentration of some pollutant in the air one breathes (that is, the individual's health status may be okay, but the material input of the oxygen producing process is in some manner inadequate or unusable). Of course, the coughing could also be a result of both inferior quality body capital and inferior air quality. In any case, it is likely that

$$St = St\ (K_t,\ Q_t,\ M_t^s) \qquad (10)$$

or that the occurence of some symptom is dependent on the true state of the individual's health, air quality, and possibly on medicines used to alleviate the advent of a symptom (i.e., one could use cough drops to reduce coughing, eye drops to reduce eye irritation, or aspirin to relieve a headache). Given this it is likely that

$$'K' \quad 'Q_t+' \quad {}_{iM}^{S}s > 0_{\text{"}} \tag{11}$$

These symptoms are the only observable manner in which the individual may get a perception of his true health state. If there are no symptoms to the contrary an individual is likely to assume he is okay while if some are prevalent he is likely to assume that something is not right. Another way in which he may evaluate his health status is to procure medical information. For example, although a person with high blood-pressure rarely has noticeable symptoms, a blood pressure test could reveal the problem and thus, give the individual a clearer picture of his health status. Also, going back to the example of coughing above, a medical check-up could tell the individual if in fact the coughing was due to something like emphysema or instead just by "something in the air" meaning his health state was okay. This suggests that

$$H_t = H_t\,(St,\ M_t^I) \tag{12}$$

or that the individual's perceived health status depends on the symptoms he observes and any additional medical information he has purchased concerning how to evaluate these symptoms or discovering health problems without current or may assume he is okay and that there is merely "something in the air" depending on his opinion and that of any medical person. In either case, his behavior will be based on his perception of his health status whether or not this perception is right or wrong. That is, an individual behaves according to the perceived state of his health and not the actual or true state. Mathematically, the individual's health problem may be stated in continuous terms as follows:

$$\max_o \int^T U(Q,\ C,\ B,\ E,\ H,\ X)e^{-rt}dt \tag{13}$$

subject to:

$$\dot{K} = g(k,\ q,\ C,\ B,\ E,\ X,\ M^K) - \delta K$$

$$s = S(K,\ Q, M^S)$$

$$H = H(S,\ M^I)$$

$$\bar{Y} \geq P_X X + P_C C + P_B B + PEE + P_M(M^K + M^S + M^1)\ \forall t$$

$$K(0) = K_0$$

where $\bar{Y}$ is the individual's income constraint and P are the various prices of the respective marketed goods. This is an optimal control problem wherein the individual's health state and his consumption of other commodities act as control variables and his true health state, K, is the state variable with its equation of motion. In other words, the individual's problem involves manipulating C, B, E, H, and X subject to a budget constraint in order to maximize his utility. A solution to this model will depend on what assumptions

are made (is $U_{Q} \gtrless 0$?) but the important tradeoffs will be adequately represented. Further note that the model allows for all three manners in which a change in air quality might effect the utility of an individual: (1) directly through aesthetic effects, (2) indirectly through changes in his body capital which will effect his health status and finally, (3) indirectly through changes in the symptoms he may observe which again effect his perception of his health status.

An important step towards the solution of this model involves the link between air quality, cigarettes, etc. and the advent of symptoms or an estimation of the symptom function, S_t. This is a primary objective of the remainder of this study.

Unfortunately, a thorough search of the medical literature has revealed practically no applicable equations to estimate even a "proxy" for health status or "body capital," or for the oxygen production function. In consequence, we have had to abandon this modelling approach and apply a more simple model structure.

Outline of the Model Applied

It has been proposed in many economic studies of health effects that individuals derive disutility from perceived and/or actual occurrences of disease. However, most individuals cannot correctly diagnose their own diseases except for a small set of common ailments. The individual commonly perceives one or more symptoms of the potential occurence of a disease. The individual may then select three alternatives, to seek medical services for diagnosis and cure; to use self-prescribed medication or other forms of self-help, or to do nothing. Typically, the individual will make these choices based on the severity of symptoms and the cost of medical services. If the symptoms are common types, i.e., the sudden appearance of a slight chest pain, the individual is likely to do nothing. Also, if the cost of medical services is extremely low or negative, the individual is likely to seek medical attention for the appearance of any symptom. The important point is that individuals work with symptoms and not the actual disease itself, whether it is the afflicted party or the physician making the diagnosis. Thus, we postulate a simple welfare relationship where S denotes a vector of symptoms and I a vector of other goods and services the individual purchases. Then the individual's utility can be represented as:

$$u = U(x, s) \tag{14}$$

where, for illustrative purposes, the function u(*) is assumed to be continuous in I and S and twice differentiable. The individual is assumed to be constrained by a budget constraint on purchases of medical services to alleviate symptoms or cure diseases and purchases of other goods and services:

$$P_X X + P_M M \leq Y \tag{15}$$

where M is the quantity of medical services, Y is income, and P denotes the unit price of the service X either as a scalar or vector. finally, to

complete this simple model, we denote a relationship between the incidence and severity of symptoms and required medical services. For simplification, it is assumed there are a fixed set of medical services to alleviate symptoms or treat various diseases, provided the individual seeks treatment and that this relationship can be expressed as:

$$M = h(S) \tag{16}$$

Next it is presumed the individual maximizes utility subject to the budget constraint and medical technologies. The first order conditions become:

$$\left.\begin{aligned} {}'_X - \lambda P_X &\leq 0 \\ {}'_S + \delta h_S &\leq 0 \\ -\lambda P_M - \delta &\leq 0 \end{aligned}\right\} \quad {}'_S + \lambda h_S \leq 0 \tag{17}$$

with $\lambda \geq 0$, $\delta \geq 0$, $u_X \geq 0$, and $u_S \leq 0$. These conditions simply indicate that the maximizing individual will purchase goods and services up to the point where marginal utility for goods is equated with the utility adjusted price of the goods. The individual will purchase a reduction in symptoms (improvement in health) up to the point where marginal disutility associated with symptoms is equal to utility adjusted productivity of purchases of medical services. Note that this follows regardless of whether there is a correct doagnosis of symptoms. What is important to the individual is whether the symptoms are alleviated and a return to good health status is perceived. A derived demand relationship for M can be developed from the presence of symptoms as follows:

$$M = f(P_X, PM, s) \tag{18}$$

where $f(\cdot)$ evolves from the first order conditions in (17). Following Mäler (1974), compensating and equivalent variation measures of consumer surplus can be constructed for S where the individual cannot control the appearance of symptoms except through changes in lifestyle or preventative actions which will not be considered here. While conceptually willingness to pay to avoid symptoms or associated medical expenses can be derived, no attempt is made in this study to estimate equation (18). The reason for not doing so is that no adequate data exist for the NAS twins to estimate M or PM. As an alternative, average U.S. medical expenditures for each type of illness were used to estimate a minimum willingness to pay to avoid symptoms . The underlying assumption is that individuals, at minimum, would be willing to pay to avoid symptoms what they typically do pay to alleviate them. In this sense, a minimum estimate is calculated.

THE DATA SET

NATIONAL ACADEMY OF SCIENCES TWIN REGISTRY*

The data which 'this research analyzes to discover the net effects of air pollutants was obtained from the NAS-NRC Twin Registry (4). This twin panel consists of 7,960 white male twin pairs, of which 6,741 twin pairs or less are examined in this study. Table 1 summarizes the age distribution of the NAS Twin panel in 1967 when the panel was asked to complete the epidemiological questionnaire (Q2) which provides the relevant health data. The twins ranged from 41 to 51 years of age at the time the Q2 information was collected. The average age was 45.

The sample itself is the result of a detailed procedure by which the National Research Council identified white male twins born during the period 1917 to 1927 in the continental United States. Additional screening was done on this set of twins to determine the twin pairs for which both members served in the armed forces (5). The process resulted in the 7,960 twin pairs currently comprising the Twin Registry.

An initial questionnaire (Q1) was used to obtain each individual's medical history since separation from military service and to identify the brothers zygosity (6). Figure 3 presents the question used on Q1 to obtain each individual's medical history since military separation. This information provides the basis for a diagnostic index which is maintained for the NAS-NRC Twin Registry. This Q1 information has been updated and purged from the diagnostic index as more complete information in medical history was collected based on Veterans Administration (VA) claims records, VA hospital records, and death certificates. In fact, the present diagnostic index is largely based on such VA information sources rather than the self-reported information from Q1.

The reader might find it tempting to consider using information in the diagnostic index to quantitatively define health status in the sort of statistical exercise which is summarized below. However, the diagnostic index represents an amalgam of different data sources each of which would be expected to contribute its own unique biases to such an analysis. For example, the self-reported Q1 information is purged when VA information is available. Therefore, the entire set of VA criteria determines the set of Q1 information that remains. Fundamentally, the VA criteria relate to military causes of medical problems as well as a certain socio-economic status. Actual information in the diagnostic index, because it is collected from different sources, may be inconsistent and therefore potential introduction of biases is difficult if not impossible to sort out.

TABLE 6.1 AGE DISTRIBUTION OF NATIONAL ACADEMY OF SCIENCES TWIN SAMPLE - 1967

Age	Absolute Frequency	Relative Frequency	Cumulative Frequency
41	1622	12.0%	12.0%
42	1646	12.2	24.2
43	1470	10.9	35.1
44	1536	11.4	46.5
45	1419	10.5	57.1
46	1265	9.4	66.4
47	1282	9.5	76.0
48	1180	8.8	84.7
49	786	5.8	90.5
50	744	5.5	96.1
51	532	3.9	100.0
TOTAL	13,482	100.0	

List **any** illness, impairment, disability, hospitalization, and operation you have had since separation from military service, stating the year when it first occurred.

Illness, impairment or operation	Year it began	**Name of Hospital**	City **and** State

Figure 6.3 **NAS Twins (Ql) Self-Reported Medical** History **Questionnaire**

And Now Some **Rather Specific Questions About Where You Have Lived Since the Second World War**

50. For consecutive periods, fill in length of period, city or community, as well as state.
Check also at the right of Table in what type of area you were living and working, respectively.

PERIOD OF TIME	CITY OR TOWN	STATE	LIVING IN			WORKING IN		
			Downtown Area	Suburban Area	Rural Area	Downtown Area	Suburban Area	Rural Area
1945 -			☐	☐	☐	☐	☐	☐
			☐	☐	☐	☐	☐	☐
			☐	☐	☐	☐	☐	☐
			☐	☐	☐	☐	☐	☐
			☐	☐	☐	☐	☐	☐
			☐	☐	☐	☐	☐	☐
			☐	☐	☐	☐	☐	☐
			☐	☐	☐	☐	☐	☐

Figure 6.4 **NAS Twins (Q2) Residence and Work History**

The epidemiological information obtained in 1967 from Q2 is the basis for the quantitative measures of health status that are utilized in the statistical analysis which is summarized here. The Q2 health status information is separated into information on respiratory and cardiovascular health problems.

Information on respiratory health status is provided by answers to two questions: do you get short of breath walking with other people at an ordinary pace on the level? Do you regularly or for extended periods of time have a cough? Clearly the binary answers to these questions are either yes or no.

With respect to cardiovascular health status a series of three binary questions provide relevant information. Have you ever had any pain or discomfort in your chest? Have you ever had a severe pain across the front of your chest lasting for a half hour or more? Have you ever had a heart attack?

The statistical analysis summarized later uses the answers to these five questions as binary dependent variables in a regression analysis. Q2 also provided information on a number of potentially relevant explanatory variables. The individual is asked by Q2 to report if he has ever had asthma, his height and weight, whether he has to diet to keep his weight down, the number of cigarettes and cigars smoked per day, as well as the individual's alcohol consumption. In addition, Q2 collects relatively detailed information on dietary habits.

A particularly interesting set of information collected by Q2 is a detailed residence and work history by location. Figure 4 presents the question used to gather this information. This type of information may be particularly useful to a statistical analysis examining the association between air pollution and human health not only because it identifies past residences by city and state, but also because it identifies if the residence and work location were in a "downtown", "suburban", or "rural" area.

Finally, a third questionnaire (Q3) collected economic information such as household income. Unfortunately, Q3 was completed by the panel in 1973 rather than 1967 when the Q2 health information was obtained. Yet Q3 provides the only economic information and 1973 household income is used as a proxy for the same 1967 variable in the statistical analysis. The actual income question was: "How much was your family income from all sources (during 1973)?"

Q3 also provided information on an individual's access to medical care. Q3 asks a detailed set of questions relating to whether the individual does or does not have an annual medical check-up. If SO, additional information is gathered on the source of payment of check-up: government clinic, union clinic, company clinic, or medical insurance.